

Rebuttal Letter

HIV self-testing positivity rate and linkage to confirmatory testing and care: a telephone survey in Côte d'Ivoire, Mali, and Senegal

Summary

I thank the authors for speedy and thorough revision of this manuscript, which I found to be of much improved quality but still in need of substantial revisions.

Below are some remaining concerns that need to be addressed by the authors before I can make my recommendation. Apologies for the level of detail, but I think it will help to make sure the revised version is much closer to a recommendation.

We sincerely thank you for the time and effort you have devoted to reviewing our manuscript. Your very detailed and precise comments helped us to refine and improve our work.

We have carefully considered all the points raised and made revisions accordingly. Please find below point-by-point responses to your comments.

Comments

Please change the keywords to the following: AIDS; HIV; Self-Testing; Key Populations; MSM; sex-workers; phone-based survey; West Africa; confirmatory testing; follow-up care; public health program evaluation.

Thank you for your comments. We have made the changes you recommended.

Line 34: This still says “reactive” - did phase two include those who said “reactive” but who had just one line? And doesn't it include those with two lines but who said “not reactive”? It's still not clear? I thought this should be “positive” - either two lines, reactive, or both. No?.

We have corrected this error. The sentence has been modified as follows:

Line 34-35: *“In the second phase (September-October), participants who had reported two lines or who reported a reactive result were recontacted to complete another questionnaire.”*

Regarding the use of “reactive” instead of “positive”, please see our detailed answer to your comment regarding Line 131.

Line 36. I think the use of the word “consistent results” may be confusing. In my opinion, it should be “consistent response”. This last option is ideal for what appears to be a set of questions used to address self-consistency in reporting. It should be stated somewhere in the methods!! Otherwise, “consistent results” has a connotation that implies consistency between multiple results, whereas what the authors mean here is coherence, or agreement (“concordance”) or “consistency” between the result (1 or 2 lines) and its interpretation (reactive or non-reactive), which is response to questions and not results of a test. So, for this line in the abstract, one might say “X% reported an HIVST interpretation (reactive vs non-reactive) that was consistent with the test results (2 lines vs 1 line, respectively)”. Obviously, if it is changed here, it should be changed throughout the manuscript. Even in Table 1 the authors use the term “consistent answer”. So I think it really should be “consistent response”. And also change in the tables and figures as needed.

We opted for “consistent response” as suggested and updated the text accordingly as well as tables and figures.

Line 38: “depending on calculations” is not specific enough. Either choose one figure and report on that, or say, “depending on whether or not incoherent results were included...”. Or “between 2.4% and 9.1%, varying by age groups, countries, response coherence, etc.” Whatever is true.

We have provided more details on the variation of the positivity rates. The sentence has been modified as follows:

Line 38-40: *“Overall HIV positivity rate was 2.5% (central hypothesis, low: 2.4%, high: 9.1%) based on self-interpreted results, and 4.5% (4.4% to 7.2%) based on the reported number of lines. Variations were observed according to country, distribution channel, sex, and age group.”*

Lines 46-47: What does this mean? Linkage was suboptimal because only 1% of original study participants responded? I think the authors mean that confirmatory testing following a positive HIVST result was low, but what is optimal? This is subjective. Please clarify. Address also in line 415 (concluding paragraph of discussion).

We have provided explanatory information by amending the last paragraphs of the summary and conclusion as follows:

Line 49-50: “Linkage to confirmatory testing following a reactive HIVST remained relatively low in these first years of HIVST implementation.”

Line 453-455: “ATLAS’ HIVST distribution strategy successfully reached people living with HIV in West Africa, although linkage to confirmatory testing following a reactive HIVST remained relatively low in these first years of HIVST implementation, and sub-optimal in the perspective of achieving UNAIDS 95-95-95 targets.”

Line 68: It’s —> It is. We try not to use contractions in formal writing. Or better yet, marry to the previous sentence. “Individuals..... result, and is therefor widely accepted.... “.

Thank you for the remark. We have corrected it.

Line 73 should read: “The HIV Self-Testing in Africa (STAR) initiative carried out in ... <https://www.psi.org/fr/project/star/>). “ (Spelling out the acronym and also moving the link up from where it currently is a few lines down).

We have modified the sentence to reflect your comment. It now reads as follows:

Line 75-76: “The HIV Self-Testing in Africa (STAR) carried in Eastern and Southern Africa and funded by Unitaïd aimed to boost the global market for HIVST (<https://www.psi.org/fr/project/star/>).”

Line 78: “where HIV epidemics differ, ...” But also: HOW do they differ? More concentrated how? Also need a reference to this.

We have modified the sentence and added a new reference:

Line 78-83: *“Following the experience gained in Eastern and Southern Africa under the STAR project [11, 23–28], the Unitaid funding agency sought to stimulate HIVST in West Africa, where HIV epidemics are distinguished by their more concentrated and less generalised nature compared to those in Eastern and Southern Africa. In this region, the general population prevalences are relatively very low, and key populations (for example, female sex workers and men who have sex with men) are particularly affected and bear a disproportionate share of the HIV burden [29].”*

Line 81: type-o delete “were”

We have corrected it.

Line 90: period (.) at end of sentence

We have added (.) at the end of the sentence.

Line 93: This sentence should have been worded as authors stated in their response to reviewers: “...second phase....conducted among those with an HIVST reactive result or had reported two lines in the first survey”. I would go one step further and say “conducted among those with an HIVST positive result or interpretation (reporting either reactive, two lines, or both during the first survey).” It’s stated differently in different parts of the manuscript, and is confusing.

The ATLAS project followed WHO wording in their HIVST guidelines, using “reactive” / “non-reactive” to qualify HIVST results, instead of “positive” / “negative”. All ATLAS documentation, brochures and videos used “reactive” and “non-reactive” wording. See also our detailed answer regarding your comment for Line 131 for the addition made to the manuscript to clarify this choice.

We have changed the sentence as follows:

Line 97-99: *“In the second phase (September-October), participants who had reported two lines or a self-interpreted HIVST result as reactive were recontacted to complete another questionnaire.”*

Line 99: Probably best to follow a comma-delimited number convention. 397 367 → 397,367

We've standardized this across the board, using the comma as a thousand separator.

Line 102: "...no previous experience with HIVST".

We have corrected it.

Lines 116-119: This reference should be inserted as a normal reference. " Figure 1. ATLAS delivery channels (adapted from [29]). FSW=

This has been corrected.

Line 131: Please edit as follows —> "should be interpreted as follows: "reactive" if two lines (C and T) are visible (even barely), "nonreactive" if only the C (control) line is visible, and "invalid" if no line is visible or if only the T (test) line is visible."

Thank you for the suggestion. We have integrated it.

Line 131: While the manufacturer's instructions instruct the tester to interpret the results as either "positive" or "negative" (or invalid), the survey responses appeared to be "reactive" or "non-reactive" (or invalid) results. Is this just a translation thing? I understand the academic jargon is "reactive" or "non-reactive", but this is hardly as clear to the general public as "positive" or "negative". Maybe it would be best to share a copy of the survey, at least in English, to understand what could be improved to limit inconsistent/incoherent results? I expected to see it in the data management plan, but either missed it or it wasn't there.

The ATLAS project followed WHO wording in their HIVST guidelines, using "reactive" / "non-reactive" to qualify HIVST result, instead of "positive" / "negative". All ATLAS documentation, brochures and videos used "reactive" and "non-reactive" wording.

It is now explained explicitly within the manuscript:

Line 131-139: *"Only oral self-testing (OraQuick HIV Self-Test®) has been distributed through ATLAS. OraSure Technologies, the manufacturer of the OraQuick test, accompanies each HIVST kit with a user manual for result interpretation (Figure 2). OraQuick HIVST results should be interpreted as*

follow: “reactive” (“positive”) if two lines (C & T) are visible (even barely), “non-reactive” (“negative”) if only the C (control) line is visible, and “invalid” if no line is visible or if only the T (test) line is visible. To be noted, the French version of the HIVST instructions distributed by ATLAS used the wording “reactive” / “non-reactive” instead of “positive” / “negative” to qualify the HIVST result, following WHO vocabulary in their HIVST guidelines[20] as an HIVST is triage test and does not provide a definitive HIV-positive diagnosis. The questionnaire of the survey also used “reactive” / “non-reactive” wording (<https://doi.org/10.5281/zenodo.10210464>).

We updated the title of Figure 2:

Line 142-145: “Figure 2. Guidelines for interpreting HIVST result, extracted from the English version of the manufacturer instructions for use (OraQuick HIV Self-Test®). To be noted, the French version distributed by ATLAS was using the wording “Reactive” / “Non-reactive” instead of “Positive” / “Negative” to qualify the HIVST result.”

To be noted, the original version of the manuscript included the French version of the instructions, but it was rejected by medXriv (all figures must be in English).

We also added examples of videos developed by ATLAS and using the wording “reactive”:

Line 147-150: “In addition to the manufacturer’s instructions, locally adapted brochures and explanatory videos in French and local languages have been developed to help users perform the test, interpret the result and know what actions should be taken following a non-reactive, a reactive or indeterminate result (for example : <https://youtu.be/laCCjVEKZto> in English or <https://youtu.be/1xzitLD309U> in French).”

Line 140: needs space between “HIVST” and “and”.

We have corrected it.

Line 155: A link to the survey given should be included here, and explicit details about the data collected. (Judging from the rest of the materials, this should be: “gender (identified as man or woman), age group (24 years or less, 25-34 years, 35 years or more), marital status (single, living with partner/married, divorced/separated/widowed...), education level (none/primary, secondary, higher...), and whether or not this was the first time they had [?? used an HIVST, been tested for HIV, ... whatever the authors mean by “first time tester”]. As a rule, any data on which analyses or results rely should be explicitly stated here.

We have added the article related to the phase 1 survey which was recently published.

Description of phase 1 questionnaire has been updated:

Line 182-186: *“The phase 1 questionnaire, which lasted 20 to 30 minutes, collected information on sociodemographic characteristics of HIVST users (including age, sex, marital status, education level), testing history (having ever tested for HIV before using HIVST and date of last HIV test), sexual and preventive behaviours, HIVST use, and difficulties encountered [39]. Specifically, each participant was asked about the number of lines that appeared when reading the HIVST result and their self-interpretation of it (reactive or non-reactive).”*

A better description of phase 2 questionnaire has been added:

Line 195-198: *“Phase 2 questionnaire asked the participants if they had undergone a confirmatory test following their HIVST, reasons for not linking to confirmatory testing (if not linked), time between HIVST and confirmatory testing (if linked), type of facility for confirmatory testing, confirmation test result, linkage to antiretroviral treatment (if confirmed HIV-positive).”*

Line 210-211: *“Both phase 1 and phase 2 questionnaires have been made available online and a link is now provided (<https://doi.org/10.5281/zenodo.10210464>).”*

The variables used in the analysis are now described in the Data analysis section:

Line 213-216: *“Following a previously published analysis [39], and due to small numbers of participants in certain distribution channels, distribution*

channels (Figure 1) were grouped in three categories: FSW-based channels (outreach or facility-based), MSM-based channels (outreach or facility-based) and other channels (PWUD-based channels, index testing, STI consultations).”

Line 229-235: “We deemed it important to stratify the positivity rates by country, sex, distribution channel, and age group (15-24, 25-34, and 35+). As the key population profile of participants should differ substantially by distribution channel (women from the FSW-based channel are more likely FSW while those from the MSM-based channel are more likely female partners of MSM; men from the MSM-based channel are more likely MSM while those from the FSW-based channel are more likely partners or clients of FSW, see Figure 1), we decided to combine distribution channel and sex into a single combined variable”.

Line 237-242: “To assess any participation bias, characteristics of phase 2 participants (country, sex and distribution channel, age group, marital status, educational level, and first-time testers, i.e. whether they ever tested for HIV before using HIVST) were compared with individuals eligible for phase 2 but who did not participate and with phase 1 participants not eligible for phase 2. Bivariate comparison was done using chi-squared tests, and multivariate comparison using a multinomial logistic regression model and then computing likelihood-ratio tests.”

Line 157: move ref 37 into the sentence (before the period(.))...

We have corrected it.

Lines 159-162: Somewhere here the authors need to state that they advertised the financial incentive on the flyer.

We have changed the paragraph as follows:

Line 164-166: “Following its results, we decided to introduce a reward as a token of appreciation for the time participants dedicated to the survey. We mentioned this financial incentive on the survey leaflet to inform the participants. Consequently, completion of the questionnaire was rewarded

with 2,000 XOF (approximately 3.40 USD) of phone communication credit. This reward was mentioned on the survey flyers.”

Line 175: The authors reference a poster here that supposedly reports on these same data. A poster, in my opinion, is not appropriate here as a reference. In addition, it draws further scrutiny because the numbers do not match. I would very much like to know the total number of kits distributed FOR THIS STUDY, as the poster provides (44,598). I also want to know why the poster says there was a total of 2405 participants while the manuscript claims 2615 participants??:

Thank you for your comment. We have replaced that reference with our other article, recently published and related to phase 1 of the survey. The number of participants is consistent between this reference and this paper.

The poster was initially limited to channels focusing on MSM, FSW, and PWD, which reduced the number of surveyed participants to 2,405.

Line 183: “completed the phase 2 questionnaire.” Also, what specific data did the phase 2 questionnaire collect? The questionnaire needs to be provided, as with phase 1. Were responses multiple choice or open-ended, for instance?)

See previous answer (Line 182-186, Line 195-198). A better description of phase 2 questionnaire has been added and both phase 1 and phase 2 questionnaires are now available online.

Lines 195-199: The wording of this paragraph/sentence could be better: “...we distinguish between those for which the reported number of visible lines was consistent with the reported self-interpretation (2 visible lines reported as reactive, one line reported as non-reactive, or no/one line and interpreted as invalid), those for which the test results were inconsistent with the interpretation, or those who returned only a partial response (refusal to answer or answered “I don’t know” to one or both questions).

Thank you for the suggestion. We have taken this into account and modified the paragraph as follows:

Line 217-222: *“Based on phase 1 participants’ self-reports, we distinguish between those who provided consistent response, i.e. the reported number of visible lines was consistent with the reported self-interpretation (2 visible lines reported as reactive, one line reported as non-reactive, or no/one line*

and interpreted as invalid), those who provided inconsistent response, i.e. the number of visible lines was inconsistent with the self-interpretation of the result, or those who returned only a partial response (refusal to answer or answered “I don’t know” to one or both questions).”

Lines 200-204: I think here, the three hypotheses would be much more clear if they are presented slightly differently, with the “central” hypothesis being the base calculation for positivity (the authors use it for the category tests later on), then saying that because of inconsistent interpretations and the eligibility of those with positive results or interpretation for phase two, they also calculated the lowest possible positivity (assuming all DK-R responses as non-reactive), and the highest possible positivity (assuming all DK-R responses as reactive) of the population surveyed. This allows the reader to understand why one would make such outrageous assumptions (neither of which is likely to be true.)

The paragraph has been rewritten:

Line 223-226: *“Due to inconsistent responses, we separately considered the self-interpreted results and the reported number of lines on the HIVST to estimate HIVST positivity rates. For each source, we excluded “don’t know” and refusals (DK-R) from both the numerator and the denominator (central hypothesis). We also computed the highest positivity (assuming all DK-R as reactive) and lowest positivity (assuming all DK-R as non-reactive).”*

Paragraph 206-209 (participation bias) should go before the positivity rates (after Line 199).

Here we do not report on bias in participating in phase 1 survey, but on participation in phase 2 survey alone.

Positivity rates are computed using only phase 1 data. And eligibility for phase 2 is based on reported HIVST test result. This is why we report on positivity rate before exploring biases in phase 2 participation.

Lines 206-209 (participation bias): “... sociodemographic and key population source characteristics were... using... ??” In Table S1 results, the authors perform multiple chi-sq tests on non-independent data without correcting for multiple tests. The more correct test here would be a log-likelihood ratio test with all response variables as arguments in an additive multinomial multiple regression model (e.g., VGAM package `vglm(response group ~ country + sex_and_distribution_channel + age group +`

marital_status + education + first_time_tester, family = "multinomial"), giving the test statistic for each variable or a results table for model comparisons.

In addition to bivariate analysis, we used multivariate analysis for comparison. All of this has been described in the text as follows:

Line 236-242: *"We described the selection of eligible participants for phase 2 questionnaires and corresponding participation rates. To assess any participation bias, characteristics of phase 2 participants (country, sex and distribution channel, age group, marital status, educational level, and first-time testers, i.e. whether they ever tested for HIV before using HIVST) were compared with individuals eligible for phase 2 but who did not participate and with phase 1 participants not eligible for phase 2. Bivariate comparison was done using chi-squared tests, and multivariate comparison using a multinomial logistic regression model and then computing likelihood-ratio tests."*

Also, authors need to explain why they group gender and distribution channels? I don't see the reason for this, since men and women can both be reached by all channels - it implies some hidden analysis that has not been made explicit. If there's an a-priori reason for separating the PWUD vs different outreach facilities, please detail it. In my opinion, it would be better to separate Men vs Women ("gender") and then facility vs community outreach-based? Then if the authors want to test which outreach-based facilities differed, they could run a posthoc test..

See previous answer.

An entry by Key Population is more relevant than opposing "facility-based" vs "outreach" as it is more aligned with the type of partners for implementation.

Line 213-216: *"Following a previously published analysis [39], and due to small numbers of participants in certain distribution channels, distribution channels (Figure 1) were grouped in three categories: FSW-based channels (outreach or facility-based), MSM-based channels (outreach or facility-based) and other channels (PWUD-based channels, index testing, STI consultations)."*

Line 231-235: *"As the key population profile of participants should differ substantially by distribution channel (women from the FSW-based channel*

are more likely FSW while those from the MSM-based channel are more likely female partners of MSM; men from the MSM-based channel are more likely MSM while those from the FSW-based channel are more likely partners or clients of FSW, see Figure 1), we decided to combine distribution channel and sex into a single combined variable.”

Line 205 - This is still not clear. Why? Just for description? For testing? If testing, with what statistical test? And the authors do not mention age class here, either. I think the best thing to do is to run the same test as for the participation bias (although a simple binomial multiple regression (glm) will suffice in this case), using only the central hypothesis. Then, depending on those results, it would make more sense how the results of the 3 hypotheses are ‘stratified’ (displayed in the figure = is it more important to stratify based on age? Country? Sex and channel?). Authors are at liberty to include results as they feel are important or were intended a priori. However, it currently feels very random as to which combinations of results are being displayed.

We conducted two binomial logistic regression analyses to examine the positivity rate in our central scenario, based on self-interpreted results on one hand, and based on the reported number of lines on the other. Although the model did not show statistical significance for variables such as age group or Sex and distribution channel, we deemed it relevant to stratify the obtained positivity rates by age group and also by the Sex and distribution channel. The detailed models have been included in the appendix for reference (Table S1).

The paragraph has been rewritten:

Line 227-230: *“We conducted two binomial logistic regression analyses to examine the positivity rate in our central scenario, based on self-interpreted results on one hand, and based on the reported number of lines on the other. We deemed it important to stratify the positivity rates by country, sex, distribution channel, and age group (15-24, 25-34, and 35+)”.*

Line 220: (prop.test function in the ‘stats’ package)

We have integrated it.

Line 238 - this should still be a part of the prior paragraph.

We have corrected it.

Line 245 Table 1: Partial answer formula “P7” needs subscripting.

Thank you for this remark. We have corrected it.

Lines 259, 332: As the third reviewer mentioned, the categories here do not match with the rest of the manuscript. The authors do not seem to have understood this comment. So, explicitly, is the lower age class 15-24, or 24 years and under? Is it “24 and under” or “under 24”? And how does discussion of “25 and older” were respondents under 15 years old excluded?

We are indeed talking about individuals aged 15-24 years. We have corrected this in the text. The concerned sentence (L259) has been modified as follows:

“When analysing positivity rates by age group (Table S2), for participants aged between 15 to 24 years old, the rates ranged from 2.2% to 7.4% based on the reported self-interpreted result and from 3.1% to 5.9% based on the reported number of lines.”

Figure 3, Tables S1, S2, etc: It is not always clear what distribution channels are being referenced. There are Community-based MSM and FSW channels, but also facility based MSM and FSW channels, and then PWUD community-based channels, and then another whole list of facility-based channels. Please be specific/explicit.

We added a paragraph in the Methods to explain it:

“Following a previous published analysis [39], and due to small numbers of participants in certain distribution channels, distribution channels (Figure 1) were grouped in three categories: FSW-based channels (outreach or facility-based), MSM-based channels (outreach or facility-based) and other channels (PWUD-based channels, index testing, STI consultations).”

Titles of Figure 3, Table S1 and Table S2 have also been updated.

Line 318: why “secondary” distribution? Figure 1 shows it was primary and secondary...?

It was an error of wording. It has been updated:

Line 354-357: *“Our study shows that the strategy implemented by the ATLAS program, through primary and secondary distribution of HIVST kits and dedicated channels, achieved HIV positivity rates of 2.5% (central hypothesis, low: 2.4%, high: 9.1%) based on self-interpreted results, and 4.5% (central hypothesis, low: 4.4%, high: 7.2%) based on the reported number of lines.”*

Lines 319 -320: Be sure to report here only the central hypothesis, unless explicitly stating how the lower and upper numbers were attained (because they are not valid results, but lower and upper limits).

We have emphasised the positivity rates derived from our central hypothesis while also considering the alternative scenarios. The paragraph has been revised as follows:

Line 354-357: *“Our study shows that the strategy implemented by the ATLAS program, through primary and secondary distribution of HIVST kits and dedicated channels, achieved HIV positivity rates of 2.5% (central hypothesis, low: 2.4%, high: 9.1%) based on self-interpreted results, and 4.5% (central hypothesis, low: 4.4%, high: 7.2%) based on the reported number of lines.”*

Lines 327-335: Here, again, just present the central hypothesis (which is a valid figure). The lower and upper limits can then be discussed as relevant (e.g., this is high, even if we had considered all missing responses to have been negative).

We have emphasised the positivity rates derived from our central hypothesis while also considering the alternative scenarios. The paragraph has been revised as follows:

Line 363-372: *“According to our estimates, HIVST positivity rates were 2.0% (central hypothesis, low: 1.8%, high: 9.8%) based on self-interpreted results, and 3.9% (central hypothesis, low: 3.8%, high: 5.4%) based on the reported number of lines in Côte d’Ivoire. In Mali, these rates were 3.6% (central hypothesis, low: 3.5%, high: 6.7%) and 5% (central hypothesis, low: 4.9%, high: 7.8%), while in Senegal, they were 1.4% (central hypothesis, low: 1.2%, high: 15%) and 6.0% (central hypothesis, low: 5.4%, high: 14.9%). Overall, these results for HIVST positivity are generally higher than the average overall positivity of HIV testing services (excluding HIVST) in West Africa. For instance, in 2020 an estimated 1.9% of all HIV tests performed were found to be positive in the region (95% credible intervals: 1.3 to 2.7%) [42]. Further, among 15-24 and 25-34 years old, which constitute more than 80% of our sample, overall positivity was 0.9% (0.7 to 1.3%) and 1.6% (1.2 to 2.2%), respectively. Collectively, these results provide evidence that HIVST is a high-yield testing modality that can address the unmet HIV testing needs of key populations and their partners.”*

Lines 341-342: How do your results suggest this? Were you able to correct your estimates for treatment? If not, how is this part of the discussion relevant. It's missing something.

We have reworded and moved the paragraphs (line 327-335 and line 337-351) as follows:

Line 363-391: *"According to our estimates, HIVST positivity rates were 2.0% (central hypothesis, low: 1.8%, high: 9.8%) based on self-interpreted results, and 3.9% (central hypothesis, low: 3.8%, high: 5.4%) based on the reported number of lines in Côte d'Ivoire. In Mali, these rates were 3.6% (central hypothesis, low: 3.5%, high: 6.7%) and 5% (central hypothesis, low: 4.9%, high: 7.8%), while in Senegal, they were 1.4% (central hypothesis, low: 1.2%, high: 15%) and 6.0% (central hypothesis, low: 5.4%, high: 14.9%). Overall, these results for HIVST positivity are generally higher than the average overall positivity of HIV testing services (excluding HIVST) in West Africa. For instance, in 2020 an estimated 1.9% of all HIV tests performed were found to be positive in the region (95% credible intervals: 1.3 to 2.7%) [42]. Further, among 15-24 and 25-34 years old, which constitute more than 80% of our sample, overall positivity was 0.9% (0.7 to 1.3%) and 1.6% (1.2 to 2.2%), respectively. Collectively, these results provide evidence that HIVST is a high-yield testing modality that can address the unmet HIV testing needs of key populations and their partners.*

Our results are in line with data collected by ATLAS implementing partners. Between 2020 and 2021, these ATLAS partners collected spontaneous feedback from HIVST users. This unpublished data collection was non-systematic and varied from one partner to another. Among 4 463 documented feedback, HIVST was reactive for 188 cases (4.2%), consistent with our estimates based on the reported number of visible lines (4.5%).

In 2021, a study based on the UNAIDS-supported Shiny90 mathematical model [43] estimated, using data from 184 population surveys and reports from national HIV screening programs from 40 sub-Saharan African countries, that the positivity rates for conventional HIV testing were 1.4% in Côte d'Ivoire, 2.2% in Mali, and 1.0% in Senegal. Our estimates for HIVST were higher than these estimates for conventional testing.

It is important to interpret HIV positivity rates while considering the treatment-adjusted prevalence (i.e., removing those on treatment from the numerator and denominator of HIV prevalence), a more reliable indicator for evaluating the effectiveness and positivity rates of targeted screening programs [44]. In West Africa, the treatment-adjusted prevalence remained relatively low in 2021: 0.6% in Côte d'Ivoire, 0.7% in Mali, and 0.06% in Senegal, according to UNAIDS data (<https://aidsinfo.unaids.org/>). Our positivity rates in each country are higher than the treatment-adjusted prevalence, suggesting that the ATLAS HIVST distribution strategy successfully reached a hard-to-reach population and at positivity levels at least as high as with passive surveillance."

Lines 342-345: Reference [43]'s modified explanation does not answer the concern here. I've read through and it appears this is an inappropriate reference for the figures cited. There is no mention of the figures stated, nor the methods on 184 population surveys. This reference is interesting because it talks about what could be expected if key populations were targeted, so maybe the authors meant to reference it elsewhere? The authors should double-check this and adjust accordingly. The methods we wanted to see were not the modeling methods ("shiny90" doesn't actually tell us anything) but the "conventional HIV testing" methods referenced. I think it's important to have these figures of "traditional" testing, because those should be fairly high - when people have a reason to get tested, there is typically a bias towards positivity ("passive surveillance"), because it tells you that your active surveillance method was identifying a hard-to-reach population and at positivity levels at least as high as with passive surveillance. This is really the crux of how you will make the case that this program was effective.

Thanks for the comment. It is indeed a mistake. We have corrected it with the correct reference:

"Maheu-Giroux et al., « National HIV Testing and Diagnosis Coverage in Sub-Saharan Africa: a new modelling tool for estimating the 'first 90' from program and survey data ». AIDS 33():p S255-S269, December 15, 2019. | DOI: 10.1097/QAD.0000000000002386."

Line 347: in line with.

We have corrected it.

Lines 353-355: How can 90% interpret correctly, but only 2% interpret incorrectly? And why is it those who reported two visible lines that suggests a problem with interpretation? Please doublecheck and re-word to clarify.

We have modified the paragraph as follows:

Line 393-396: *“In our study, 2.0% of the participants reported an inconsistent response between the number of visible lines and their self-interpretation of the result and 6.0% reported a number of lines but didn’t know how to interpret it or refused to answer, suggesting potential issues in interpreting the number of visible lines on HIVST kits.”*

Lines 368-372: I would add here that the instructions given with the test (Figure 2) give conflicting instructions. They say to interpret as “Positive” or “Negative”, instead of “Reactive” and “Non-reactive”. For instance, hearing two different terms used, I would have thought that 2 lines (C + / T +) = reactive and positive, one line (C + / T -) = reactive and negative (as in, the control worked, it reacted), and one or no line (C - / T x) = non-reactive/invalid. Throughout this manuscript, I was under the distinct impression that the use of the term “reactive” or “non-reactive” was purely academic for reporting here, and not a colloquial usage for the surveys. This highlights why it is so important for the surveys - in the languages delivered to participants - need to be shared here, so that further improvements can be made.

See our previous answer.

The ATLAS project followed WHO wording in their HIVST guidelines, using “reactive” / “non-reactive” to qualify HIVST result, instead of “positive” / “negative”. All ATLAS documentation, brochures and videos used “reactive” and “non-reactive” wording.

It is now explained explicitly within the manuscript:

Line 131-139: *“Only oral self-testing (OraQuick HIV Self-Test®) has been distributed through ATLAS. OraSure Technologies, the manufacturer of the OraQuick test, accompanies each HIVST kit with a user manual for result interpretation (Figure 2). OraQuick HIVST results should be interpreted as follow: “reactive” (“positive”) if two lines (C & T) are visible (even barely), “non-reactive” (“negative”) if only the C (control) line is visible, and “invalid” if no line is visible or if only the T (test) line is visible. To be noted, the French version of the HIVST instructions distributed by ATLAS used the wording “reactive” / “non-reactive” instead of “positive” / “negative” to qualify the*

HIVST result, following WHO vocabulary in their HIVST guidelines[20] as an HIVST is triage test and does not provide a definitive HIV-positive diagnosis. The questionnaire of the survey also used “reactive” / “non-reactive” wording (<https://doi.org/10.5281/zenodo.10210464>)”

We updated the title of Figure 2:

Line 142-145: *“Figure 2. Guidelines for interpreting HIVST result, extracted from the English version of the manufacturer instructions for use (OraQuick HIV Self-Test®). To be noted, the French version distributed by ATLAS was using the wording “Reactive” / “Non-reactive” instead of “Positive” / “Negative” to qualify the HIVST result.”*

To be noted, the original version of the manuscript included the French version of the instructions, but it was rejected by medXriv (all figures must be in English).

We also added examples of videos developed by ATLAS and using the wording “reactive”:

Line 147-150: *“In addition to the manufacturer’s instructions, locally adapted brochures and explanatory videos in French and local languages have been developed to help users perform the test, interpret the result and know what actions should be taken following a non-reactive, a reactive or indeterminate result (for example : <https://youtu.be/laCCjVEKZto> in English or <https://youtu.be/1xzitLD309U> in French).”*

Lines 374-383: State first the results, then compare to other estimates as you have done, then conclude with the first sentence at the end of the paragraph to say that while this study’s numbers were low, they still showed a lower rate of linkage to confirmatory testing. I would maybe also provide the linkage by country, because a country’s system for availability of traditional testing facilities may have a role to play, or in the case of this study, due to the language used the interviewer (which was the clear leader of reasons why confirmatory testing was not sought)

We have been more explicit in modifying the paragraph as follows:

Line 413-419: *“Linkage to confirmatory testing following a reactive test was 44% (95% confidence interval from 33% to 55%). However, this estimate*

includes some individuals who did not adequately self-interpreted their HIVST result as reactive. When considering only those who reported two lines and self-interpreted their result as reactive, the linkage rate increased to 56% (36% to 74%). This percentage is closer to that was observed in a study conducted in Kenya on HIV testing of FSW male partners using HIVST secondary distribution, where 65% of men with a reactive result had a confirmatory test [47]. Our estimates were based on small numbers resulting in large confidence intervals, but are still showing a low rate.”

Line 403: the cost - not necessary, but do the authors have any information on the cost of this study? In reference to the overall cost of HIVST programs?

We do not have a precise estimate of the total costs of the survey as, in addition to the direct costs, it would require including all indirect human resources costs.

Lines 407-408: Again, remove this reference to a poster which gives conflicting results, and could be considered as a prior publication of these results, which would mean we would have to reject this article. Please explain in your response (not in the article) why this poster’s analysis showed different numbers and also be sure that all authors from that poster are included in this peer-reviewed submission.

This reference has been changed by the article recently published on phase 1 of the survey.

Line 416 : remove the term “reactive”. Maybe say “a positive self-test result” ? Again, be consistent with the rest of the data/methods.

We have provided explanations in our previous answers, which justify the use of the terms "reactive" and "non-reactive".

Lines 426 (Table S1): Two important questions: 1) Why were there 3 people with 2 lines and reactive, non-reactive, or DK-R for reactivity in the “Not eligible for Phase 2) column?? Does “Not eligible” mean they also did not consent to be recontacted? If so, please change the language of this column. 2) Because eligibility is dependent on the results of the test, authors should not be testing “Result and number line” differences between the columns. Instead, I think “results consistency” could be

tested here, because the column does not depend on this (either 2 lines or a “reactive” response qualifies for eligibility, if I understood correctly)?

Thank you for your comment. It was indeed an error. The individuals concerned have been reclassified in the column 'eligible not having participated in phase 2 of the survey'. The table has been modified accordingly. We have also added the global p-values for the multivariate analysis.

Lines 439 (Table S4) : the percentages on the column-totals should not be there. They don't make sense with the rest of the table. **Check to make sure there is discussion about how a respondent could have given consistent responses but said their reason for no follow-up test was because their test was non-reactive...

We have removed the total column.

Just a suggestion: if available, Efficacy could be estimated by understanding the % of new infections from the countries reported for 2021 that came from the STAR/ATLAS program...?

We believe that an accurate estimation of the effectiveness of the ATLAS program would be feasible with more information on new infections detected during the implementation period of the project in these three countries. A comparison of the rates of new infections before and after the project, supported by other data sources, would be necessary to achieve this.

To be noted, there is ongoing modelling work to estimate the epidemiological impact of ATLAS and of scaling-up ATLAS strategy:
<https://www.medrxiv.org/content/10.1101/2023.08.23.23294498v2>.

Some parameters of the model were estimated based on our phone survey.

Table S7: Apologies - I originally said this comment was for table S5. In fact, it is for Table S7: please arrange from shortest to longest time periods.

We have reclassified the periods from shortest to longest.

Please check on how the term “sex” was used. Did the survey use the terms “man” and “woman” for “sex” (which should be “male” and “female”, as man and woman typically refer to genders). And also comment on whether there was any ambiguity about e.g., transgender identity.

The questionnaire used the wording “Quel était votre sexe ? Homme ou Femme”. It should be noted that the nuance male/female vs man/woman in English doesn’t translate exactly in French.

The wording used in French is much more related to sex, and would not be relevant to collect self-identified gender, where a question as “Vous définiriez-vous plutôt comme un homme ou une femme ?” would be more appropriate in such a context.

The survey was not designed to collect transgender identities and collecting such identities in West Africa is complex. Questionnaires used for LGBTIQ+ surveys in Europe or North America are often not directly applicable in West Africa as they miss nuances about gender roles. See for example this preprint about Cote d’Ivoire : <https://osf.io/preprints/socarxiv/xmfpn/>